NEW YORK UNIVERSITY
INSTITUTE OF [illegible] CAL SCIENCES
25 W[illegible] 3, N. Y.

EM 96

AFCRC-TN-56-592
ASTIA DOCUMENT NO. AD 98791

NEW YORK UNIVERSITY
Institute of Mathematical Sciences
Division of Electromagnetic Research

RESEARCH REPORT No. EM-96

Diffraction by an Aperture.II

J. B. KELLER, R. M. LEWIS and B. D. SECKLER

CONTRACT NO. AF 19(122)-42

AUGUST 1956

EM-96

NEW YORK UNIVERSITY

Institute of Mathematical Sciences

Division of Electromagnetic Research

Research Report No. EM-96

DIFFRACTION BY AN APERTURE. II

Joseph B. Keller, Robert M. Lewis and Bernard D. Seckler

Joseph B. Keller
Joseph B. Keller

Robert M. Lewis
Robert M. Lewis

Bernard Seckler
Bernard D. Seckler

Morris Kline
Morris Kline
Project Director

August, 1956

The research reported in this document has been made possible through the support and sponsorship extended by the Air Force Cambridge Research Center, under Contract No. AF 19(122)42. It is published for technical information only and does not necessarily represent recommendations or conclusions of the sponsoring agency.

New York, 1956

3cp

Abstract

The field diffracted by an aperture of any shape is evaluated asymptotically for small wavelengths on the basis of several approximate diffraction theories. These are the Kirchhoff method, the two customary modifications of it, and W. Braunbek's new modification of this method. In each case a double integral over the aperture is evaluated asymptotically, and contributions from interior stationary points, edge stationary points, and corners of the edge are obtained. The contributions from points of each type, according to all the theories examined, are exactly of the form recently deduced by J. B. Keller using his geometrical theory of diffraction. The edge stationary points and the corners correspond respectively to rays singly diffracted from edges and from corners. For edge diffracted rays the Braunbek diffraction coefficients alone coincide with those given by the geometrical theory of diffraction, but the Braunbek method does not apply to corner diffraction. None of these methods takes account of multiple diffraction as does the geometrical theory of diffraction.

Table of Contents

1. Introduction 1

2. The Kirchhoff and Braunbek methods 3

3. Asymptotic evaluation of the Kirchhoff integrals 6

4. Asymptotic evaluation of the Braunbek integrals 15

5. The transmission cross section 22

6. Conclusion 27

Appendix I Interpretation of u^I 30

Appendix II The geometrical factor G 31

Appendix III Asymptotic evaluation of Rubinowicz' line integral 33

References 36

1. Introduction

Diffraction by an aperture in a thin screen was treated by a new method in Part I[1] of this paper. In this second part we compare this method with certain other methods which have been used in the past to treat the same problem. The other methods which we consider are the Kirchhoff method, the two customary modifications of this method[2] and W. Braunbek's new modification of it[3]. These methods have been carefully discussed by C. J. Bouwkamp[2].

The geometrical theory of diffraction used in Part I involves the concept of a diffracted ray. It enables us to express the singly diffracted field explicitly in terms of geometrical quantities. Similar geometrical expressions can also be obtained for the multiply diffracted field. In the Kirchhoff method and its modifications, on the other hand, the diffracted field is expressed as a double integral over the aperture. We will evaluate the double integrals given by all these methods asymptotically for small wavelengths λ. We will find that each one leads to the formula deduced in Part I for the singly edge-diffracted field, with an appropriate diffraction coefficient in each case. This fact supports the theory used in Part I. Only the Braunbek method yields exactly the same edge diffraction coefficient, and therefore the same singly edge-diffracted field, as is obtained in Part I. This agreement between the Braunbek method and the geometrical theory of diffraction lends support to both of these methods, and shows that the simpler geometrical theory can be used instead of the more complicated Braunbek method.

Neither the Braunbek method nor the Kirchhoff methods take account of multiple diffraction, which is included in the theory of Part I. Furthermore, the Braunbek method cannot be applied to apertures with corners. The other Kirchhoff methods can be applied to such apertures and yield results of the form given in Part I for the field on a ray diffracted from a corner. However the Kirchhoff diffraction coefficients differ from those of Part I, which come from exact solutions of certain diffraction problems involving corners. Therefore the Kirchhoff diffraction coefficients are inaccurate, although they are easy to compute. It is possible to modify the Braunbek method so that it applies to apertures with corners and yields results for them in agreement with those of Part I. But even then it will not take account of multiple diffraction.

When the transmission cross section of an aperture is computed by any of the Kirchhoff methods using the cross section theorem, only the geometrical optics cross section is obtained. The Braunbek method yields a correction of order k^{-2} to the geometrical cross section. However the geometrical theory of diffraction yields a correction of order $k^{-3/2}$ for a circular hole in a screen on which $u=0$. The failure of the Kirchhoff and Braunbek methods to yield the proper correction in this case can be attributed to their failure to account for multiple diffraction.

In spite of the defects of the Braunbek and Kirchhoff methods pointed out above, they do have one advantage over the method of Part I. That method yields infinite values on caustics of the diffracted rays and on shadow boundaries. It must be supplemented by special considerations in order to yield finite results at such places. The Braunbek and Kirchhoff methods do not have this defect. Therefore they can be used to supplement the method of Part I on the surfaces where it fails.

For the special case of an incident spherical wave, A. Rubinowicz[4] has evaluated the original Kirchhoff integral asymptotically. He did it by transforming the double integral into a line integral along the rim of the aperture and then evaluating the line integral by stationary phase. His result, which he interpreted in terms of diffracted rays, is a special case of the result which we will obtain. The integrals given by the modified forms of the Kirchhoff method were evaluated asymptotically by N.G. van Kampen[5]. His results are the same as those we obtain in these cases and his heuristic method of evaluation has been justified by J. Focke[6] and N. Chako[7]. However neither of their analyses is directly applicable to the integrals which occur in the Braunbek method. These integrals are our main concern, since they have not previously been evaluated asymptotically. We have treated them by a method different from those given in [5], [6], [7]. For the sake of uniformity we have also treated the modified Kirchhoff integrals by this same method.

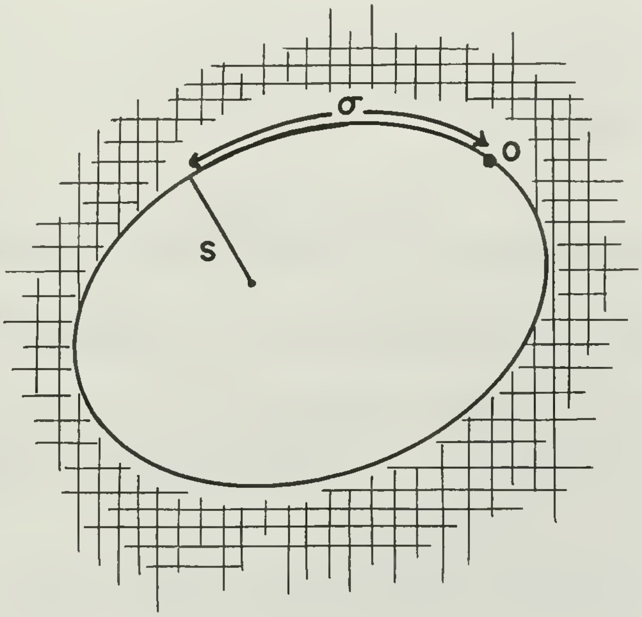

Figure 1
The coordinates σ and s of a point in the aperture. Distance along the edge from some fixed point is denoted by σ. The σ coordinate of a point is the value of σ at the nearest point on the edge. The s coordinate is the distance from the point to the edge.

2. The Kirchhoff and Braunbek methods

Consider a thin screen S containing an aperture H and lying in the plane $z = 0$. Suppose an incident scalar wave u_{inc} comes from the left $(z < 0)$ and strikes the screen. We wish to determine the resulting field u, assuming that it satisfies the reduced wave equation. The total field u must consist of the incident field u_{inc}, to the left of the screen, and an additional scattered or outgoing part on both sides of the screen. Furthermore on the screen either u or its normal derivative must vanish, depending upon the type of screen. Thus u satisfies one of the boundary conditions

$$u = 0 \qquad \text{on } S \tag{1}$$

$$\frac{\partial u}{\partial z} = 0 \qquad \text{on } S. \tag{2}$$

By means of Green's theorem, the field u behind the screen $(z > 0)$ can be represented as an integral over the plane $z = 0$ in three different ways:

$$u(P) = \frac{1}{4\pi} \int \left[u \frac{\partial}{\partial z} \left(\frac{e^{ikR}}{R} \right) - \frac{e^{ikR}}{R} \frac{\partial u}{\partial z} \right] dxdy, \tag{3}$$

$$u(P) = \frac{1}{2\pi} \int u \frac{\partial}{\partial z} \left(\frac{e^{ikR}}{R} \right) dxdy, \tag{4}$$

$$u(P) = -\frac{1}{2\pi} \int \frac{e^{ikR}}{R} \frac{\partial u}{\partial z} \, dxdy. \tag{5}$$

In these equations R denotes the distance from the point P, at which u is evaluated, to a point on the plane $z = 0$. In the integrands, u and $\frac{\partial u}{\partial z}$ are to be evaluated on the back side $(z = 0+)$ of the screen. The propagation constant $k = 2\pi/\lambda$, where λ is the wavelength.

In order to compute u(P) from the first of these formulas we must know u and $\frac{\partial u}{\partial z}$ on the plane z = 0. We know that one or the other of these quantities is zero on S according as boundary condition (1) or (2) is satisfied. Kirchhoff assumed that both u and $\frac{\partial u}{\partial z}$ are zero on the rear of S and that $u = u_{inc}$, $\frac{\partial u}{\partial z} = \frac{\partial u_{inc}}{\partial z}$ in the aperture H. Then (3) becomes

$$(6) \qquad u_K(P) = \frac{1}{4\pi}\int_H \left[u_{inc}\ \frac{\partial}{\partial z}\left(\frac{e^{ikR}}{R}\right) - \frac{e^{ikR}}{R}\ \frac{\partial u_{inc}}{\partial z}\right] dxdy.$$

We will call the field given by this method the Kirchhoff approximation and denote it by $u_K(P)$.

If (4) is used instead of (3), only the value of u need be known. Suppose that (1) is satisfied. Then, if we assume that $u = u_{inc}$ in H, (4) becomes

$$(7) \qquad u_1(P) = \frac{1}{2\pi}\int_H u_{inc}\ \frac{\partial}{\partial z}\left(\frac{e^{ikR}}{R}\right) dxdy.$$

We will call the field given by (7) the first modified Kirchhoff approximation and denote it by $u_1(P)$. In a similar way, if (5) is used and (2) is satisfied, the assumption $\frac{\partial u}{\partial z} = \frac{\partial u_{inc}}{\partial z}$ in H leads to

$$(8) \qquad u_2(P) = -\frac{1}{2\pi}\int_H \frac{e^{ikR}}{R}\ \frac{\partial u_{inc}}{\partial z}\ dxdy.$$

We will call the field given by (8) the second modified Kirchhoff approximation and denote it by $u_2(P)$.

The three fields u_K, u_1 and u_2 are all approximate descriptions of the diffracted field since they are based upon approximate values of u and/or $\frac{\partial u}{\partial z}$ on the plane z = 0. Physically it seems clear that the assumed values on this plane should be the more nearly correct the smaller the wavelength λ or the larger the value of k. Therefore we will evaluate the integrals which represent these fields asymptotically for k large. Before doing so it is convenient to notice that

$$u_K = \frac{1}{2}(u_1 + u_2). \tag{9}$$

Therefore we need only evaluate u_1 and u_2, and then we can find u_K from (9). The evaluations of the integrals are contained in the next sections.

The new modification of the Kirchhoff method proposed by Braunbek makes use of the representation (4) when $u = 0$ on S, and of (5) when $\frac{\partial u}{\partial z} = 0$ on S. In the case $u = 0$ on S it assumes that $u = u_{inc}$ in H except near the edge. Near any point Q on the edge u is assumed to be given by the field near a half-plane on which $u = 0$. The half-plane lies in the plane of the screen, and its edge is tangent to the edge of the screen at Q. A plane wave is incident on it from the direction, and with the amplitude, of the field incident on the aperture at Q. Since the field diffracted by a half-plane is known, an explicit expression for the field near any point Q on the edge is available. In the case $\frac{\partial u}{\partial z} = 0$ on S it is assumed that $\frac{\partial u}{\partial z} = \frac{\partial u_{inc}}{\partial z}$ in H except near the edge. Near the edge $\frac{\partial u}{\partial z}$ has the value given by the field near a half-plane on which $\frac{\partial u}{\partial z} = 0$. We will call the two fields given by (4) and (5), with the Braunbek integrands, u_{B1} and u_{B2} respectively.

Just as in the case of the Kirchhoff approximation and its usual modifications, it is to be expected that the Braunbek approximation becomes more accurate as k increases. Therefore we will also evaluate the integrals given by this approximation for large k. It also seems likely that this approximation is more accurate than the previous ones. However it is clear that the Braunbek method does not apply near corners of the edge. The obvious modification required at such corners is that the field there should be replaced by the field near an 'infinite corner,' i.e. a screen with two straight edges meeting at a corner.

In evaluating the Kirchhoff and Braunbek integrals we will assume that the incident field has the form

$$u_{inc} \sim Ae^{ik\Psi} \tag{10}$$

The phase Ψ and amplitude A are functions of position. Since a field of this form is not necessarily a solution of the reduced wave equation, we have used the sign '$\sim$' in (10). This indicates that the incident field is only asymptotically of this form, for large k.

3. Asymptotic evaluation of the Kirchhoff integrals

Upon inserting the incident field (10) into (7) and (8) we obtain

$$u_1(P) \sim \frac{ik}{2\pi}\int_H AR^{-1}R_z\, e^{ik(R+\Psi)}dxdy, \tag{11}$$

$$u_2(P) \sim -\frac{ik}{2\pi}\int_H AR^{-1}\Psi_z\, e^{ik(R+\Psi)}dxdy. \tag{12}$$

Note that in both equations only the leading terms in k have been retained.

When an integral such as that in (11) or (12) is evaluated asymptotically for k large, the entire contribution to the asymptotic value comes from the neighborhoods of certain critical points[6],[7]. These points are of three kinds. Critical points of the first kind, which may be called interior stationary points, are interior points of H at which the phase $R+\Psi$ is stationary. Critical points of the second kind, which may be called edge stationary points, are points of the edge at which the phase is stationary with respect to variations along the edge. Critical points of the third kind, or corners, are, to the order we consider, points of the edge at which the slope is discontinuous. Let us designate the contributions of the three kinds of stationary points to $u_1(P)$ by $u_1^I(P)$, $u_1^{II}(P)$ and $u_1^{III}(P)$ with a similar notation for $u_2(P)$. Then we have

$$(13) \qquad u_1(P) \sim u_1^I(P) + u_1^{II}(P) + u_1^{III}(P),$$

$$(14) \qquad u_2(P) \sim u_2^I(P) + u_2^{II}(P) + u_2^{III}(P).$$

At an interior stationary point $(R+\Psi)_x = (R+\Psi)_y = 0$. Since $(\nabla\Psi)^2 = (\nabla R)^2 = 1$, while $\Psi_z > 0$ and $R_z < 0$, it follows that $\nabla\Psi = -\nabla R$. Thus, noting that ∇R points from P toward the aperture, we see that the incident ray in the direction $\nabla\Psi$ must, when extended, pass through P. Therefore to locate critical points of the first kind we first extend all the incident rays through the plane $z = 0$. If a ray which passes through P goes through the aperture H, then its point of intersection with H is a critical point of the first kind. Thus if P is in the geometrically lit region, there is such a stationary point, while if P is in the geometrically dark region there is none.

The contributions $u_1^I(P)$ and $u_2^I(P)$ of the neighborhoods of interior stationary points to the integrals in (11) and (12) are, up to terms of order k^{-2},

$$(15) \qquad u_1^I(P) = u_2^I(P) = \sum \frac{\varepsilon A \; \Psi_z \; e^{ik(R+\Psi)}}{\sqrt{|\Delta|}} + O\left(\frac{1}{k^2}\right).$$

Here ε and Δ are defined by

(16) $$\Delta = (R+\Psi)_{xx}(R+\Psi)_{yy} - (R+\Psi)^2_{xy} ,$$

(17) $$\varepsilon = \begin{cases} +1 & \text{if } \Delta > 0 \text{ and } (R+\Psi)_{xx} > 0 \\ -1 & \text{if } \Delta > 0 \text{ and } (R+\Psi)_{xx} < 0 \\ -i & \text{if } \Delta < 0 . \end{cases}$$

All quantities in each term in (15) are to be evaluated at an interior stationary point, and the sum extends over all of these points. Since at such points $\Psi_z = -R_z$, both (11) and (12) yield the same result. Equation (15) can be obtained from (11) or (12) by expanding the phase of the integrand in a Taylor series (up to quadratic terms) around each stationary point and then applying the usual method of stationary phase twice. It can also be obtained by making use of equation (109) of J. Focke[6]. If $\Delta = 0$ at any point, then P lies on a caustic of the incident rays and that point contributes a more complicated term to (13).

In Appendix I it is shown that (15) is exactly equal to $u_{inc}(P)$. Therefore we define $u_{geom}(P)$ as zero if P is in the shadow and as $u_{inc}(P)$ if P is in the lit region. Then we may say that

(18) $$u_1^I(P) = u_2^I(P) = u_{geom}(P).$$

If P lies on a shadow boundary the corresponding interior stationary point becomes a boundary stationary point. In this case only half the neighborhood lies in the aperture, provided that the stationary point is not at a corner. Thus the contribution of this point is just one-half of the typical term in (15). If there is only one stationary point then that point contributes $\frac{1}{2}\, u_{inc}(P)$.

Let us now consider boundary stationary points. We could obtain the contributions from such points by using Focke's equation (123). However it is simpler to evaluate these contributions directly. To this end we introduce the arclength σ along the edge and the distance s along the normal to the edge as new variables (see Figure 1). In a sufficiently narrow strip along any smooth portion of the edge each point has just one pair of coordinate values σ and s. The area element in these coordinates is $|1 - s\rho^{-1}(\sigma)|d\sigma\, ds$ where $\rho(\sigma)$ is the radius of curvature of the edge, measured positively when the center of curvature is on the same side of the edge as the aperture. In terms of these coordinates we have

$$(19) \qquad u_1^{II}(P) \sim \frac{ik}{2\pi} \int_0^L \int_0 AR^{-1}R_z \left[1 - \rho^{-1}(\sigma)s\right] e^{ik(R+\Psi)} ds\, d\sigma.$$

In (19) L denotes the length of the edge. The upper limit on the s integration has not been indicated since the only contributions to $u_1^{II}(P)$ come from the edge, which is at s = 0.

We now integrate by parts with respect to s in (19) and retain only the leading term in k. We also ignore the contribution from the upper limit in accordance with the definition of $u_1^{II}(P)$. Then (19) becomes

$$(20) \qquad u_1^{II}(P) \sim -\frac{1}{2\pi} \int_0^L AR^{-1}R_z(\Psi_s+R_s)^{-1} e^{ik(R+\Psi)} d\sigma.$$

The integral in (20) can be evaluated by stationary phase and yields

$$(21) \qquad u_1^{II}(P) \sim \sum D_1 A e^{ik(\Psi+R)} G\ .$$

The quantities G and D_1 in (21) are defined by

$$(22) \qquad G = \frac{\sin\beta}{R\sqrt{|(\Psi+R)_{\sigma\sigma}|}}\, e^{-i\pi/4 \pm i\pi/4},$$

$$(23) \qquad D_1 = \frac{-e^{i\pi/4} R_z}{\sqrt{2\pi k}\,\sin\beta(\Psi_s+R_s)}.$$

In (21) the sum extends over the edge stationary points, at each of which $\Psi_\sigma + R_\sigma = 0$. Each term in the sum is evaluated at one of these points. The angle β at each point is the angle between $\nabla\Psi$ (the incident ray direction) and the edge. The sign in the exponent of (22) is, at each point, the same as the sign of $(\Psi+R)_{\sigma\sigma}$.

The condition for a point on the edge to be an edge stationary point is $\Psi_\sigma + R_\sigma = 0$. This can be written as $\nabla\Psi\cdot t = -\nabla R\cdot t$, where t is a unit tangent to the edge. In the latter form this condition can be interpreted geometrically. It states that the incident ray makes the same angle with the edge as does the line from the edge to P. Furthermore the incident ray and the line to P lie on opposite sides of the plane normal to the edge at the stationary point. If we call the line from the edge to P a diffracted ray, this condition is just the law of diffraction[1]. Thus we see that $u_1^{II}(P)$ contains exactly one term for each edge diffracted ray through P, and no other terms.

Precisely the same evaluation applies to $u_2^{II}(P)$ if we replace $-R_z$ by Ψ_z in the integrand. Thus the final result for $u_2^{II}(P)$ is also given by (21), but with D_1 replaced by

$$(24) \qquad D_2 = \frac{e^{i\pi/4}\,\Psi_z}{\sqrt{2\pi k}\,\sin\beta(\Psi_s+R_s)}.$$

Then from (9), $u_K^{II}(P)$ is also given by (21) with D_1 replaced by

$$D_K = \frac{1}{2}(D_1+D_2). \tag{25}$$

In Appendix II it is shown that

$$G = \left[R(1+\rho_1^{-1}R)\right]^{-1/2}. \tag{26}$$

Here R denotes the distance to P along the diffracted ray from the point of diffraction and $-\rho_1$ denotes the distance to the caustic of the diffracted rays. With this value of G inserted, (21) is exactly the same as equation (17) of Part I. That equation expresses the field at P on all rays singly diffracted from the aperture edge. Therefore $u_1^{II}(P)$ may be interpreted as the contribution to $u_1(P)$ from such rays. Then D_1 is the diffraction coefficient for edge diffraction given by the first modified Kirchhoff approximation. Since $u_2^{II}(P)$ and $u_K^{II}(P)$ are also given by (21) they also correspond to the field on rays singly diffracted from the edge. D_2 and D_K are then the edge diffraction coefficients given by the second modified Kirchhoff approximation and the usual Kirchhoff approximation respectively.

The diffraction coefficients D_1, D_2, and D_K can be written in a more geometrical form in terms of the angles α, θ and β introduced in Part I. As we have already explained, β is the angle between $\nabla\Psi$ (the incident ray) and the edge. To define α and θ we project the incident and diffracted rays ($\nabla\Psi$ and $-\nabla R$) into the plane normal to the edge at the edge stationary point (the point of diffraction). The angles between these projections and the negative z-direction are α and θ respectively (see Figure 2). In terms of these angles we have

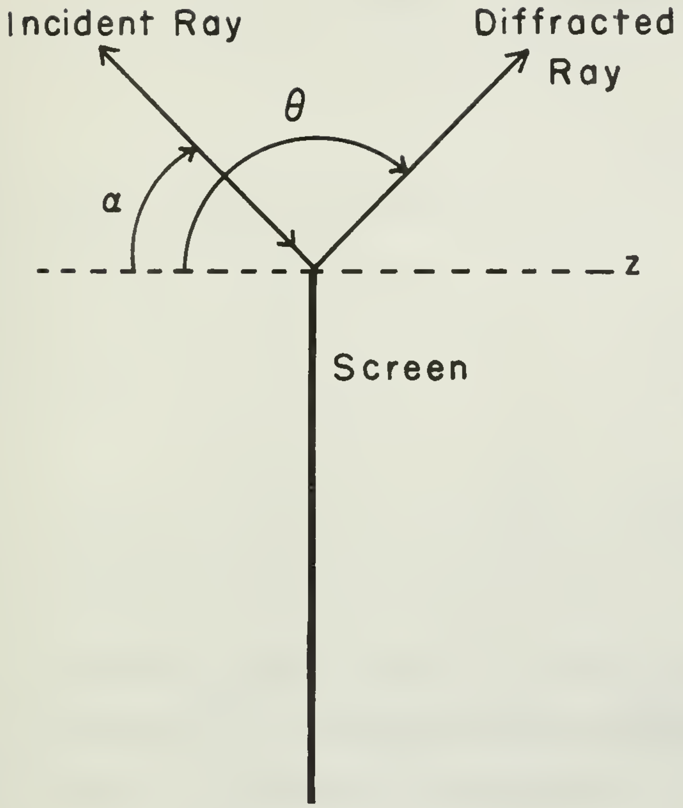

Figure 2
The projections of the incident and diffracted rays into a plane normal to the edge at the stationary point. The angles α and θ are the angles between these projections and the negative z axis, measured as shown in the figure. The edge is perpendicular to the plane of the figure.

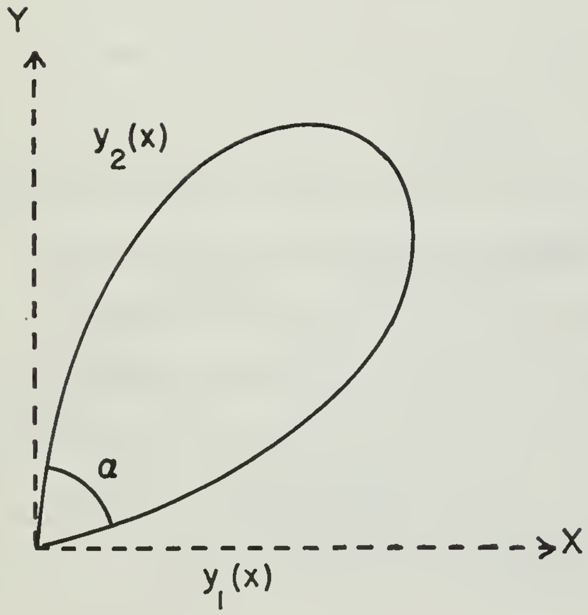

Figure 3
The x, y coordinates at a corner of the edge. The angle between the tangents to the edge at the corner is d.

$R_z = \sin\beta \cos\theta$, $Y_z = \sin\beta \cos\alpha$, $R_s = -\sin\beta \sin\theta$ and $Y_s = -\sin\beta \sin\alpha$. Inserting these expressions into (23), (24), and (25) yields

$$D_1 = \frac{e^{i\pi/4} \cos\theta}{\sqrt{2\pi k} \sin\beta(\sin\theta + \sin\alpha)}, \tag{27}$$

$$D_2 = \frac{-e^{i\pi/4} \cos\alpha}{\sqrt{2\pi k} \sin\beta(\sin\theta + \sin\alpha)}, \tag{28}$$

$$D_K = \frac{e^{i\pi/4} (\cos\theta - \cos\alpha)}{2\sqrt{2\pi k} \sin\beta(\sin\theta + \sin\alpha)}. \tag{29}$$

It is to be noted that these diffraction coefficients become infinite when $\theta = \pi + \alpha$, which occurs if P lies on the shadow boundary. This is due to the vanishing of the factor $Y_s + R_s$ in the denominators of (23),(24) and (25). When this factor vanishes, the phase at the edge stationary point is also stationary with respect to variations into the interior. The integration by parts which led to (20) is then invalid and another method of evaluating the integral (19) is required. The result of such an evaluation was referred to after equation (18) and will not be considered further.

Finally we must consider $u_1^{III}(P)$, the field contribution from the corners. At each of these points the tangent to the edge is discontinuous, and we will denote by d the angle between the tangents at a typical corner (see Figure 3). Then if the x-axis is tangent to one edge at this corner and if the origin is at the vertex, the other edge is tangent to the line $y = x \tan d$. We will write the equations of the two edges as $y = y_1(x)$ and $y = y_2(x)$, where $y_1'(0) = 0$ and $y_2'(0) = \tan d$. Now we can write $u_1^{III}(P)$ in the form

$$(30) \qquad u_1^{III}(P) \sim \frac{ik}{2\pi} \sum \int_0 \int_{y_1(x)}^{y_2(x)} AR^{-1}R_z e^{ik(R+\Psi)} dydx .$$

The sum extends over the corners of the edge and at each corner the coordinates are situated in the manner just described. The upper limit of the x integration has been omitted, since $u_1^{III}(P)$ includes only the contribution to the integral from the corner, which is at the lower limit $x = 0$.

We now integrate by parts in (30), first with respect to y and then with respect to x. Each time we retain only the leading term in k, which comes from the endpoints. In this way we obtain

$$(31) \qquad u_1^{III}(P) \sim \sum \frac{ik}{2\pi} \int_0 AR^{-1}R_z \left. \frac{e^{ik(R+\Psi)}}{ik(R+\Psi)_y} \right]_{y_1(x)}^{y_2(x)} dx$$

$$\sim \sum \frac{AR^{-1}R_z e^{ik(R+\Psi)}}{2\pi ik(R+\Psi)_y} \left[\frac{1}{\left[R(x,y_1)+\Psi(x,y_1)\right]_x} - \frac{1}{\left[R(x,y_2)+\Psi(x,y_2)\right]_x} \right] .$$

This can be rewritten in the form

$$(32) \qquad u_1^{III}(P) \sim \sum C_1 A \frac{e^{ik(\Psi+R)}}{R}$$

where

$$(33) \qquad C_1 = - \frac{i}{2\pi k} \frac{R_z}{(R+\Psi)_y} \left[\frac{1}{\left[R(x,y_1)+\Psi(x,y_1)\right]_x} - \frac{1}{\left[R(x,y_2)+\Psi(x,y_2)\right]_x} \right] .$$

In order to express c_1 in geometrical terms, we introduce the angles a and b between $\nabla\Psi$ and the two edges at the corner, and the angle c between $\nabla\Psi$ and the z-axis (see Figure 4). Similarly we define a' and b' as the angles between $-\nabla R$ and the two edges, and c' as the angle between $-\nabla R$ and the z-axis. Then $R_z = -\cos c'$, $R_x+\Psi_x = \cos a - \cos a'$, and $\left[R(x,y_2)+\Psi(x,y_2)\right]_x = \dfrac{\cos b - \cos b'}{\cos d}$. Upon using these relations in (33) we obtain

$$(34) \qquad C_1 = \frac{i}{2\pi k} \, \frac{\cos c' \sin d}{(\cos a-\cos a')(\cos b-\cos b')} \, .$$

A similar evaluation shows that $u_2^{III}(P)$ is also given by (32), with C_1 replaced by C_2. We obtain C_2 by replacing $-R_z$ by Ψ_z in (33). Thus we find

$$(35) \qquad C_2 = \frac{i}{2\pi k} \, \frac{\cos c \sin d}{(\cos a-\cos a')(\cos b-\cos b')} \, .$$

Then from (9) we see that $u_K^{III}(P)$ is given by (32) with C_1 replaced by C_K, where

$$(36) \qquad C_K = \frac{i}{4\pi k} \, \frac{(\cos c+\cos c') \sin d}{(\cos a-\cos a')(\cos b-\cos b')} \, .$$

Equation (32) for any u^{III} is identical with equation (15) of Part I which yields the field on rays singly diffracted from corners of the edge. Thus we can say that each u^{III} also represents the field on such rays, and that C_1, C_2 and C_K are the corner diffraction coefficients given by the three Kirchhoff theories. These coefficients become infinite on the cones a' = a and b' = b. When either a' = a or b' = b, an edge stationary point coincides with the corner, and a different method of evaluating the integrals is then required. Thus these cones are analogous to shadow boundaries. When both a' = a and b' = b, an

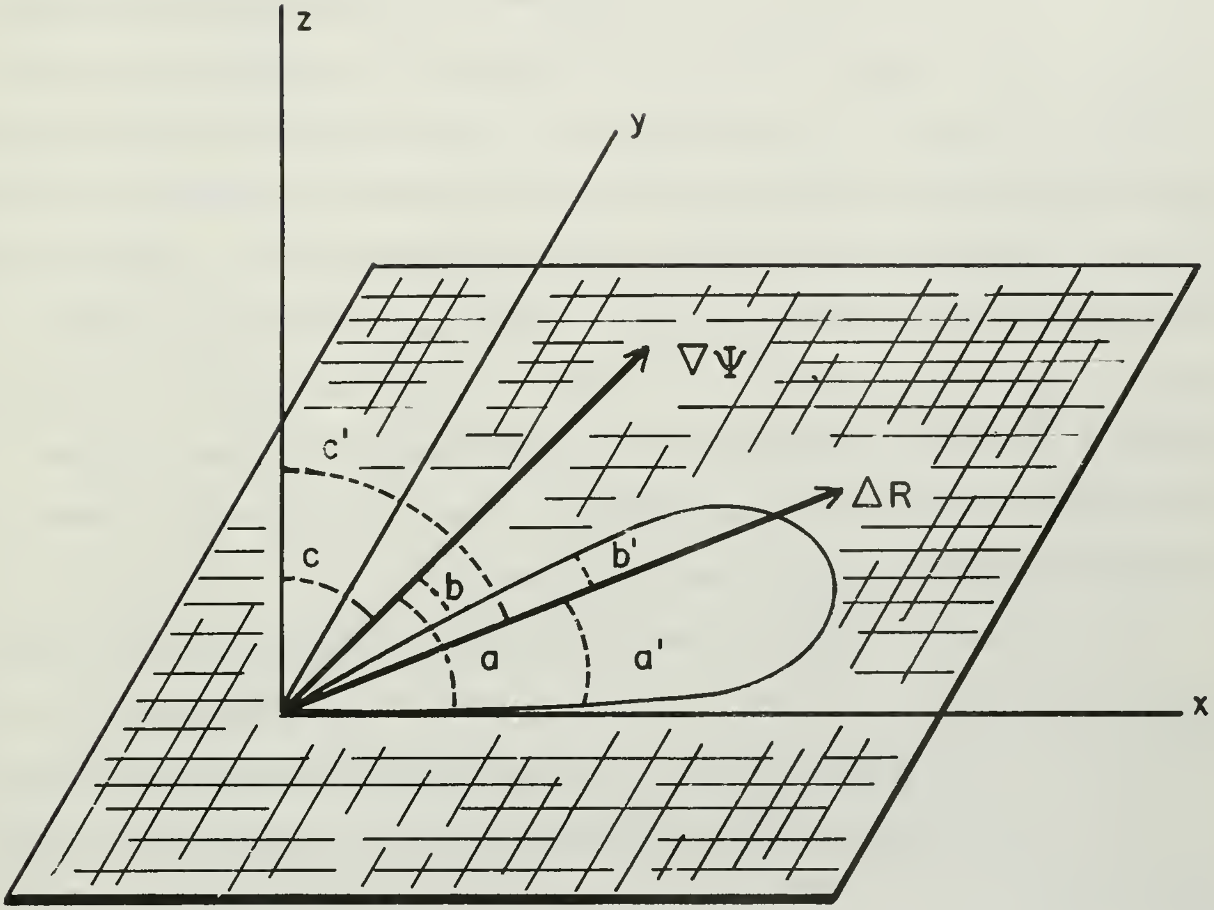

Figure 4

The incident ray and a diffracted ray at a corner of the edge. The angles between the incident ray and the two tangents at the corner are a and b; the corresponding angles for the diffracted ray are a' and b'.

'interior' stationary point coincides with the corner and a still different method of evaluation is required. We will not consider either of these cases further.

Let us summarize our results. We have seen that the field given by each of the three Kirchhoff methods is, for large k, a sum of three parts. One part is just u_{geom}, which describes the incident field on the rays which pass through the aperture. The other two parts describe, respectively, the field on the rays singly diffracted from the edge and on the rays singly diffracted from corners of the edge. These fields have exactly the form determined by the geometrical theory of diffraction. Each method yields an edge diffraction coefficient D and a corner diffraction coefficient C which differ from the coefficients of the geometrical theory. However, the diffracted field given by any one of these methods can be constructed by the geometrical theory of diffraction, if the appropriate diffraction coefficients are used.

In the special case of a spherical incident wave, u_K can be written as u_{geom} plus a line integral along the aperture edge. This was shown by Rubinowicz[4]. In Appendix III we show that the asymptotic evaluation of the line integral yields exactly $u_K^{II} + u_K^{III}$, which serves as a check on our previous calculations.

4. Asymptotic evaluation of the Braunbek integrals

The Braunbek approximations, which we have called u_{B1} and u_{B2}, are given by equations (4) and (5) respectively, with appropriate integrands. The integrands are determined by the incident field at points away from the edge, and by the solutions of appropriate half-plane diffraction problems at points near the edge. We will assume that the integrands are continuous and have several continuous derivatives. Then the asymptotic value of each integral for k large comes from the neighborhoods of the three types of critical points. Once these points are found, it will suffice to know the behavior of the integrands near them in order to evaluate the integrals asymptotically. Therefore it will not be necessary to construct explicitly the complete integrands, which join the interior and edge values together smoothly.

In the interior of the aperture the integrands of u_{B1} and u_{B2} are the same as those of u_1 and u_2 respectively. Therefore u_{B1} and u_{B2} have the same interior stationary points as u_1 and u_2 and also receive the same contribution from each point. If we designate the sum of these contributions by u_{B1}^{I} and u_{B2}^{I} then we have $u_{B1}^{I} = u_1^{I} = u_{geom}$, and $u_{B2}^{I} = u_2^{I} = u_{geom}$, since both u_1^{I} and u_2^{I} equal u_{geom}.

Let us now locate the edge stationary points and evaluate u_{B1}^{II} and u_{B2}^{II}, the contributions to u_{B1} and u_{B2} from these points. To do this we must consider the integrands near the edge. We will express them in terms of the coordinates σ and s. These coordinates are unique for points sufficiently near the edge, provided that the edge has no corners. Then, according to Braunbek, the integrand at (σ, s) is determined by diffraction from a half-plane tangent to the edge at $(\sigma, 0)$. The field incident upon the half-plane must be a plane wave. Its amplitude, phase and direction must coincide with those of the actual incident field $u_{inc} \sim Ae^{ik\Psi}$ at the point $(\sigma, 0)$ on the edge. We will write $A(\sigma)$, $\Psi(\sigma)$ and $\nabla\Psi(\sigma)$ to indicate the values of A, Ψ and $\nabla\Psi$ at this point. Then the plane wave incident upon the half-plane may be written as

$$(37) \qquad u_0(r) = A(\sigma)e^{ik\left[\Psi(\sigma)+r\cdot\nabla\Psi(\sigma)\right]}.$$

Here r denotes the position vector of a point, measured from an origin at $(\sigma, 0)$ on the edge.

Now let $u(\sigma, s)$ and $u_z(\sigma, s)$ denote the field and its normal derivative at the point (σ, s) in the aperture due to diffraction of the plane wave (37) from the half-plane tangent to the edge at $(\sigma, 0)$. Both u and u_z can be determined from the exact solution of the half-plane problem[2] originally solved by A. Sommerfeld. In using this solution, it is necessary to take account of the fact that $\nabla\Psi(\sigma)$ makes an angle β with the edge of the half-plane. Then the solution for the screen on which $u = 0$ can be evaluated in the aperture and yields

$$(38) \qquad u(\sigma, s) = \frac{2e^{-i\pi/4}}{\sqrt{\pi}} A(\sigma) e^{ik\left[\Psi(\sigma) + sn\cdot\nabla\Psi(\sigma)\right]} F_1(\sigma, s),$$

where

$$(39) \qquad F_1(\sigma, s) = \int_0^{\sqrt{2ks\sin\beta}\,\cos(\frac{\pi}{4} - \frac{\alpha}{2})} e^{i\tau^2}\, d\tau .$$

In (38) n denotes the unit normal to the edge pointing into the aperture, β is the angle between $\nabla\Psi(\sigma)$ and the edge, and α was defined previously (see Figure 2). Similarly if $u_2 = 0$ on the screen we obtain

$$(40) \qquad u_z(\sigma, s) = \frac{2e^{-i\pi/4}}{\sqrt{\pi}} A(\sigma) e^{ik\left[\Psi(\sigma) + sn\cdot\nabla\Psi(\sigma)\right]} F_2(\sigma, s)$$

where

$$(41) \qquad F_2(\sigma, s) = ik\sin\beta\cos\alpha \int_0^{\sqrt{2ks\sin\beta}\,\cos(\frac{\pi}{4} - \frac{\alpha}{2})} e^{i\tau^2}\, d\tau - \sin(\frac{\pi}{4} - \frac{\alpha}{2})\sqrt{\frac{k\sin\beta}{2s}} \cdot e^{2iks\sin\beta\cos^2(\frac{\pi}{4} - \frac{\alpha}{2})} .$$

We now insert (38) into (4) and obtain for u_{B1}^{II}

$$(42) \qquad u_{B1}^{II} \sim \frac{ke^{i\pi/4}}{\pi^{3/2}} \int_0^L \int_0 A(\sigma)\frac{R_z}{R} e^{ik\left[\Psi(\sigma)+sn\cdot\nabla\Psi(\sigma)+R\right]} F_1(\sigma,s)\left|1-\rho^{-1}(\sigma)s\right| ds\,d\sigma .$$

Similarly, inserting (40) into (5) yields for u_{B2}^{II}

$$(43) \qquad u_{B2}^{II} \sim \frac{-e^{-i\pi/4}}{\pi^{3/2}} \int_0^L \int_0 A(\sigma)\, \frac{e^{ik\left[\Psi(\sigma)+sn\cdot\nabla\Psi(\sigma)+R\right]}}{R} F_2(\sigma,s)\left|1-\rho^{-1}(\sigma)s\right| ds\,d\sigma .$$

In (42) and (43) the upper limit of the s integration has been omitted since, from the definitions of u_{B1}^{II} and u_{B2}^{II}, only contributions from the edge at the lower limit $\sigma = 0$ are to be included.

To evaluate (42) we integrate by parts with respect to s and obtain

$$(44) \qquad \int_0 \frac{R_z}{R} e^{ik\left[sn\cdot\nabla\Psi+R\right]} F_1\left|1-\rho^{-1}s\right|\, ds \sim \left.\frac{e^{ik\left[sn\cdot\nabla\Psi+R\right]}F_1\left|1-\rho^{-1}s\right|}{ikR\left[n\cdot\nabla\Psi+R_s\right]}\right|_0$$

$$- \int_0 \frac{e^{ik\left[sn\cdot\nabla\Psi+R\right]}}{ikR\left[n\cdot\nabla\Psi+R_s\right]} \; \frac{\partial F_1}{\partial s}\left|1-\rho^{-1}s\right|\, ds .$$

In the second term on the right side in (44) we have retained only the leading term in k, which comes from differentiating F_1 since $\frac{\partial F_1}{\partial s} = O(\sqrt{k})$. The terms obtained by differentiating other factors are of lower order in k since $F_1 = O(1)$, and therefore they are omitted. We see from (39) that $F_1 = 0$ at $s = 0$, so the first term on the right side of (44) is zero. Thus when $\frac{\partial F_1}{\partial s}$ is computed from (39) and inserted, (44) becomes

$$(45)\qquad \int_0 \frac{R_z}{R} e^{ik\left[sn\cdot\nabla\Psi+R\right]} F_1 \left|1-\rho^{-1}s\right| ds \sim i\sqrt{\frac{\sin\beta}{2k}}\cos\left(\frac{\pi}{4}-\frac{\alpha}{2}\right)\int_0 e^{ik\Phi}\frac{R_z\left|1-\rho^{-1}s\right| ds}{\sqrt{s}\,R(R_s+n\cdot\nabla\Psi)},$$

where

$$(46)\qquad \Phi = R + s\left[n\cdot\nabla\Psi+2\sin\beta\cos^2\left(\frac{\pi}{4}-\frac{\alpha}{2}\right)\right].$$

To evaluate the integral in (45) we employ the following asymptotic result[8],

$$(47)\qquad \int_0 e^{ikh(s)}\frac{g(s)}{\sqrt{s}}\,ds \sim \frac{\sqrt{\pi}\,g(0)}{\sqrt{kh_s(0)}}\,e^{ikh(0)+i\pi/4}.$$

Before using this in (45) we note that $\Phi_s = R_s+n\cdot\nabla\Psi+\sin\beta+\sin\beta\sin\alpha$. At $s = 0$, $R_s = \nabla R\cdot n$ and $n\cdot\nabla\Psi = -\sin\beta\sin\alpha$. Thus $\Phi_s(0) = \sin\beta+\nabla R\cdot n$. Now using (47) in (45) yields

$$(48)\qquad \int_0 \frac{R_z}{R} e^{ik\left[sn\cdot\nabla\Psi+R\right]} F_1 \left|1-\rho^{-1}s\right| ds \sim \frac{ie^{i\pi/4}}{k}\sqrt{\frac{\pi\sin\beta}{2(\sin\beta+\nabla R\cdot n)}}\;\frac{\cos\left(\frac{\pi}{4}-\frac{\alpha}{2}\right)R_z e^{ikR}}{(\nabla\Psi\cdot n+\nabla R\cdot n)R}.$$

All quantities in the right member of (48) are to be evaluated at $s = 0$, i.e. on the edge. Of course they still depend upon σ.

We now insert the right side of (48) into the right-hand member of (42), which becomes a single integral with respect to σ. This integral can be evaluated by stationary phase. In fact the phase of the integrand is just $R+\Psi$, which is exactly the same as the phase of the integral for u_1^{II} in equation (20). Therefore the stationary points are the same as those previously considered and the result of the evaluation is again of the form (21) with D_1 replaced by

$$(49)\qquad D_{B1} = \frac{e^{i\pi/4} R_z \cos(\frac{\pi}{4} - \frac{\alpha}{2})}{\sqrt{\pi k \sin\beta(\sin\beta + \nabla R\cdot n)(\nabla\psi\cdot n + \nabla R\cdot n)}} .$$

In terms of the angles previously defined we have $R_z = \sin\beta\cos\theta$, $\nabla\psi\cdot n = -\sin\beta\sin\alpha$, and $\nabla R\cdot n = -\sin\beta\sin\theta$. Upon substituting these expressions into (49) and simplifying the result, we obtain

$$(50)\qquad D_{B1} = -\frac{e^{i\pi/4}}{2\sqrt{2\pi k}\sin\beta}\left[\sec\tfrac{1}{2}(\theta - \alpha) + \csc\tfrac{1}{2}(\theta + \alpha)\right] .$$

Now u^{II}_{B1} is given by (21) with D_{B1} inserted in place of D_1. The resulting expression is exactly the same as the field given by the geometrical theory of diffraction on rays singly diffracted from edges. Therefore u^{II}_{B1} can also be interpreted as the field on the edge diffracted rays and D_{B1} as the edge diffraction coefficient given by the Braunbek method for screens on which $u = 0$. This edge diffraction coefficient is exactly the same as that of the geometrical theory for screens on which $u = 0$, which is given in Part I, equation (13).

A procedure similar to that used for u^{II}_{B1} applies to the evaluation of u^{II}_{B2}. We first insert (40) into (5) and obtain

$$(51)\qquad u^{II}_{B2} \sim \frac{-e^{-i\pi/4}}{\pi^{3/2}} \int_0^L \int_0 A(\sigma)\, \frac{e^{ik[\psi(\sigma) + sn\cdot\nabla\psi(\sigma) + R]}}{R}\, F_2(\sigma, s)\,|1 - \rho^{-1}(\sigma)s|\, ds\, d\sigma$$

Next we integrate by parts with respect to s, but we only retain the first term in $F_2(\sigma, s)$ since that is the leading term in k. In this way we get an integral of the type given in (47). Then using (47) we finally obtain

$$(52)\qquad \int_0 A(\sigma \; \frac{e^{ik|\Psi+sn\cdot\nabla\Psi+R|}}{R}\,|1-\rho^{-1}s|\,ds \sim -e^{i\pi/4}\sqrt{\frac{\pi\,\sin\beta}{2(\sin\beta+\nabla R\cdot n)}}$$

$$\left[\frac{\sin\beta\,\cos\alpha\,\cos(\frac{\pi}{4}-\frac{\alpha}{2})}{\nabla\Psi\cdot n+\nabla R\cdot n}+\sin(\frac{\pi}{4}-\frac{\alpha}{2})\right]\frac{e^{ikR}}{R}\,.$$

Now inserting (52) into (43), we express u_{B2}^{II} as a single integral with respect to σ. Evaluating this by stationary phase again leads to (21) with D_1 replaced by D_{B2} which is

$$(53)\qquad D_{B2}=\frac{e^{i\pi/4}}{\sqrt{\pi k}}\sqrt{\frac{\sin\beta}{\sin\beta+\nabla R\cdot n}}\left[\frac{\cos\alpha\,\cos(\frac{\pi}{4}-\frac{\alpha}{2})}{\nabla\Psi\cdot n+\nabla R\cdot n}+\frac{\sin(\frac{\pi}{4}-\frac{\alpha}{2})}{\sin\beta}\right].$$

In terms of the angles previously considered this becomes

$$(54)\qquad D_{B2}^{II}=-\frac{e^{i\pi/4}}{2\sqrt{2\pi k}\,\sin\beta}\left[\sec\tfrac{1}{2}(\theta-\alpha)-\csc\tfrac{1}{2}(\theta+\alpha)\right].$$

This Braunbek diffraction coefficient is exactly the same as that given by the geometrical theory of diffraction, Part I,equation (13), for a screen on which the normal derivative of the field is zero. Thus for such screens the Braunbek method also gives exactly the same field on rays singly diffracted from the edge as does the geometrical theory.

If we attempt to determine u_{B1}^{III} or u_{B2}^{III}, the contributions from the corners of the edge, we find that the Braunbek method fails. This is due to the fact that a point in the aperture near a corner is near two different points of the edge. Therefore the Braunbek method yields two different values for the integrand at such a point. Consequently the integrand near a corner is not uniquely determined, so the corner contribution cannot be evaluated. However,

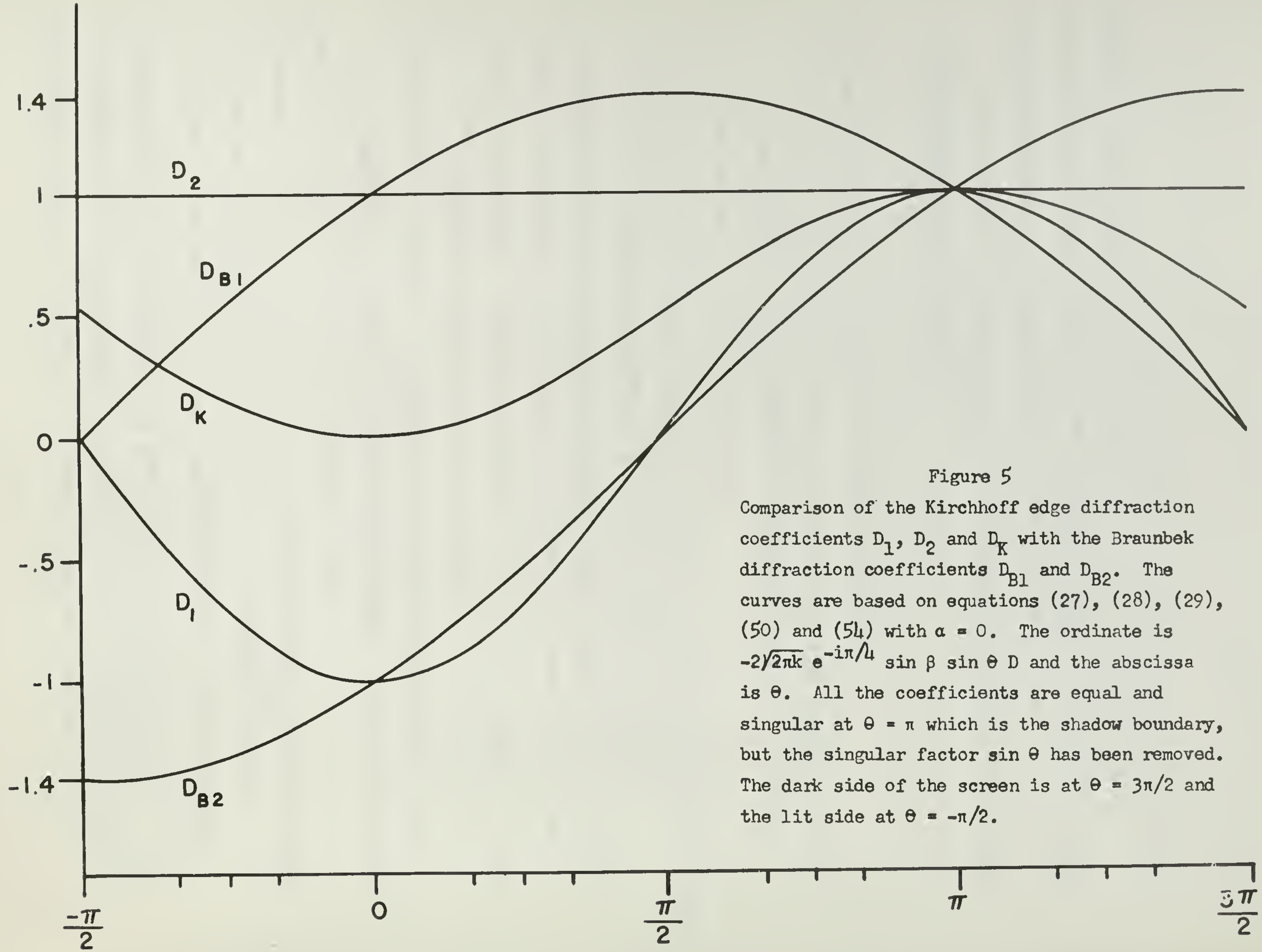

Figure 5

Comparison of the Kirchhoff edge diffraction coefficients D_1, D_2 and D_K with the Braunbek diffraction coefficients D_{B1} and D_{B2}. The curves are based on equations (27), (28), (29), (50) and (54) with $\alpha = 0$. The ordinate is $-2\sqrt{2\pi k}\, e^{-i\pi/4} \sin\beta \sin\theta\, D$ and the abscissa is θ. All the coefficients are equal and singular at $\theta = \pi$ which is the shadow boundary, but the singular factor $\sin\theta$ has been removed. The dark side of the screen is at $\theta = 3\pi/2$ and the lit side at $\theta = -\pi/2$.

the Braunbek method can be modified by assuming that the field near a corner is the same as that near an infinite corner hit by a plane wave. By an infinite corner we mean a screen bounded by two semi-infinite straight lines which meet at the corner. This modification would undoubtedly yield a field on corner diffracted rays in agreement with that given by the geometrical theory of diffraction. But this cannot be verified at present because of the intractable form of the solution of the corner diffraction problem.

In summary, we have found that the Braunbek method applies to an aperture without corners. For large k it yields a field which is a sum of two parts. One is u_{geom}, the incident field on the rays which pass through the aperture. The other is the field on rays singly diffracted from the edge. This field is of exactly the same form as that given by the geometrical theory of diffraction. For the two types of screen considered the diffraction coefficients are precisely the same as those given by the geometrical theory. These coefficients are compared with the Kirchhoff coefficients in Figure 5 . However the Braunbek method does not take account of multiple diffraction.

At caustics and on shadow boundaries the evaluations of the Braunbek integrals which we have given are invalid. At such places the integrals can be evaluated by a slight modification of the present procedure.

5. The transmission cross section

The transmission cross section σ of an aperture is defined as the ratio of the flux through the aperture to the flux per unit area in the incident wave, assuming that the incident wave is plane. According to the cross section theorem, σ can be determined from the amplitude of the transmitted field in the forward direction, i.e. in the direction of propagation of the incident wave. Let the field in this direction at a large distance R from the aperture be given by

$$(55) \qquad u \sim -\frac{ke^{ikR}}{2\pi R} f .$$

Here f is a constant. Then if the phase of the incident wave is zero at the origin of R, the cross section theorem states[2]

$$(56) \qquad \sigma = \operatorname{Im} f .$$

Of course, this theorem applies to the exact solution u. However we will apply it to the approximate solutions which we have been considering.

Let us first assume that the incident wave is $e^{ik(x \sin \gamma + z \cos \gamma)}$. Then (7) yields

$$(57) \qquad u_1(P) = \frac{1}{2\pi}\int_H \frac{e^{ik(x \sin \gamma + z \cos \gamma + R)}}{R} R_z \left(ik - \frac{1}{R}\right)dxdy .$$

If P is far from the aperture in the forward direction, the phase in (57) is essentially constant over the aperture. Furthermore R_z is practically $-\cos \gamma$ and $1/R$ is negligible compared to ik. Thus (57) yields

$$(58) \qquad u_1(P) \sim -\frac{ke^{ikR}}{2\pi R} \cdot iA \cos \gamma .$$

Here A is the area of the hole. Now (56) yields

$$(59) \qquad \sigma_1 = A \cos \gamma .$$

We see that the first modified Kirchhoff approximation yields for σ exactly the projected area of the aperture. This is just the geometrical optics cross section. This result is precise and not merely valid for large k. The second modified Kirchhoff approximation yields the same result and therefore so does the original Kirchhoff method.

If we now consider the Braunbek method, we find that the interior points of the aperture also contribute the expression in (58) to the field and $A \cos \gamma$ to the cross section. This is so because at interior points the Braunbek integrand is the same as the Kirchhoff integrands which we have just considered. However there is an additional contribution to the Braunbek field from the neighborhood of the edge. This is given in equations (42) and (43) with $A(\sigma) = 1$, $\Psi = x \sin \gamma$ and $\nabla\Psi = (\sin \gamma, 0, \cos \gamma)$. When these values are inserted the phase of the integrand becomes constant and equal to R at points far from the aperture in the forward direction. Then R_z is practically equal to $-\cos \gamma$ and R is essentially constant over the aperture. Then (42) and (43) become

$$(60) \qquad u_{B1}^{II} \sim \frac{-ke^{ikR+i\pi/4} \cos \gamma}{\pi^{3/2} R} \int_o^L \int_o F_1(\sigma, s) |1-\rho^{-1} s| ds d\sigma ,$$

$$(61) \qquad u_{B2}^{II} \sim \frac{-e^{ikR-i\pi/4}}{\pi^{3/2} R} \int_o^L \int_o F_2(\sigma, s) |1-\rho^{-1} s| ds d\sigma .$$

To evaluate the integral in (60) we first integrate by parts with respect to s, ignoring the contribution from the upper limit, and obtain

$$(62) \qquad \int_o F_1(o, s) |1-\rho^{-1} s| ds \sim -\frac{1}{4} \sqrt{kp} \int_o (2s^{1/2} - \rho^{-1} s^{3/2}) e^{ikps} ds .$$

Here we have introduced p, defined by

$$(63) \qquad p = 2 \sin \beta \cos^2\left(\frac{\pi}{4} - \frac{\alpha}{2}\right).$$

Now integrating the first term on the right of (62) by parts once, and the second term twice, we get

$$(64) \qquad \int_0 F_1(\sigma, s)\left|1-\rho^{-1}s\right| ds \sim -\frac{\sqrt{kp}}{4}\left[\frac{i}{kp} + \frac{3}{4k^2p^2\rho}\right]\int_0 \frac{e^{ikps}}{\sqrt{s}}\, ds$$

$$\sim -\frac{\sqrt{kp}}{4}\left[\frac{i}{kp} + \frac{3}{4k^2p^2\rho}\right]\sqrt{\frac{\pi}{kp}}\, e^{i\pi/4} .$$

In the second line of (64) we have inserted the asymptotic value of $\int_0 s^{-1/2}\, e^{ikps}\, ds$.

Upon inserting (64) into (60), and adding to this (58), which represents the contribution of interior points, we obtain

$$(65) \qquad u_{B1}(P) \sim -\frac{ke^{ikR}}{2\pi R}\left[iA \cos \gamma + \frac{\cos \gamma}{2k}\int_0^L \frac{d\sigma}{p} - \frac{3i \cos \gamma}{8k^2}\int_0^L \frac{d\sigma}{p^2\rho}\right] .$$

Using the cross section theorem we now find

$$(66) \qquad \sigma_{B1} \sim A \cos \gamma - \frac{3 \cos \gamma}{8k^2}\int_0^L \frac{d\sigma}{p^2\rho} .$$

In the case of normal incidence ($\gamma = 0$) (65) and (66) become

$$(67) \qquad u_{B1}(P) \sim \frac{-ke^{ikR}}{2\pi R}\left[iA + \frac{L}{2k} - \frac{3\pi i}{4k^2}\right],$$

$$(68) \qquad \sigma_{B1} \sim A - \frac{3\pi}{4k^2} .$$

The evaluation of u^{II}_{B2} is practically the same as that of u^{II}_{B1}. The integrand now involves two terms since F_2 is a sum of two terms. The first of these terms, which is the only one we retained in our previous calculations, is just $ik \sin\alpha \cos\beta\, F_1(\sigma,s)$. Since α and β depend upon σ only, this term yields the second two terms in (65), but with the factor $\cos\delta$ replaced by $\sin\alpha\cos\beta$ under the integral. Therefore in place of (65) and (66) we obtain

$$(69) \qquad u_{B2}(P) \sim \frac{-ke^{ikR}}{2\pi R}\left[iA\cos\delta + \frac{1}{2k}\int_0^L \frac{\sin\beta\cos\alpha}{p}\,d\sigma - \frac{3i}{8k^2}\int_0^L \frac{\sin\beta\cos\alpha}{p^2\rho}\,d\sigma\right],$$

$$(70) \qquad \sigma_{B2} \sim A\cos\delta - \frac{3}{8k^2}\int_0^L \frac{\sin\beta\cos\alpha}{p^2\rho}\,d\sigma .$$

At normal incidence (69) and (70) yield (67) and (68) for $u_{B2}(P)$ and σ_{B2}. When the second term in the integrand of u^{II}_{B2} is evaluated, it adds terms in $1/k$ and i/k^2 to the bracketed expression in (69) and a term in $1/k^2$ to (70).

Let us now compare the results of this section for the transmission cross section with other results. Consider first the case of normal incidence on a circular hole in a screen on which $u = 0$. In Part I it is shown that the first correction to the geometrical cross section in this case is of the order $k^{-3/2}$. This correction is obtained by taking account of doubly diffracted rays. The Kirchhoff methods yield no correction and the Braunbek method yields (68), all failing to yield the $k^{-3/2}$ correction. Next consider arbitrary incidence upon an aperture of any shape in a screen on which $u = 0$. In Part I it is shown that usually the doubly diffracted rays lead to a correction of order k^{-2} in this case, just as the Braunbek result (68) does. However the result based upon doubly diffracted rays involves the aperture size and shape, while the Braunbek method yields the same correction for all apertures. Finally let us consider normal incidence on a circular hole in a screen on which the normal derivative of u is zero. In this case H. Levine[9] has shown that the correction is $-\frac{3\pi}{8k^2}$ instead of $-\frac{3\pi}{4k^2}$, which is given by the Braunbek method in (68). When the second term in the Braunbek integrand for u_{B2}^{II} is included the disagreement is further increased.

On the basis of these comparisons we must conclude that the Braunbek method fails to yield the proper correction to the cross section when the cross section theorem is used. However the Braunbek method does yield the correct field on singly diffracted rays, since it agrees with the field given by the geometrical theory of diffraction. The disagreement in the former case and the agreement in the latter case can be explained as follows. In calculating the field on singly diffracted rays use was made of only the first two derivatives of the Sommerfeld half-plane solution along the normal to the edge, evaluated at the edge. On the other hand, in computing the cross section correction additional derivatives of the Sommerfeld half-plane solution were used. We may therefore conclude that only the first two derivatives of the aperture field at the edge are correctly given by the half-plane solution for k large. This, then, is perhaps the proper statement of the Braunbek hypothesis.

6. Conclusion

We have evaluated asymptotically, for small λ or large k, the integrals which give the field diffracted by an aperture of any shape in a thin screen.

The five integrals which we have considered are those given by the Kirchhoff method, its two usual modifications, and Braunbek's new modification for two types of screen. In every case the asymptotic form of the field is a sum of contributions from three types of critical points. These are interior stationary points, edge stationary points, and corners of the edge. All the methods lead to exactly the same critical points, and these points correspond precisely to certain rays given by the geometrical theory of diffraction. Interior stationary points correspond to rays directly transmitted through the aperture; edge stationary points correspond to rays singly diffracted at the edge, and corners correspond to rays singly diffracted from corners.

The contribution to the field from each type of point is, for every method, of exactly the form given by the geometrical theory of diffraction for the corresponding type of ray. Each method yields certain edge and corner diffraction coefficients, except the Braunbek method, which does not apply to apertures with corners. The edge diffraction coefficients given by the Braunbek method are precisely the same as those given by the geometrical theory of diffraction for both types of screen. Thus the Braunbek method yields exactly the same field as that given by the direct and singly edge diffracted rays of the geometrical theory of diffraction. It could be modified to apply to corners but it takes no account of multiple diffraction, as does the geometrical theory. The other Kirchhoff methods yield fields which correspond to direct rays, to singly edge-diffracted rays and to singly corner-diffracted rays. They yield the same field on the direct rays as do the geometrical theory and the Braunbek method. However they yield fields on the edge diffracted rays which differ from each other and from the Braunbek method, which agrees with the geometrical theory.

We conclude that the results just described lend considerable support to the geometrical theory of diffraction. They also show that, for apertures without corners, the Braunbek method is equivalent to the geometrical theory provided that only singly diffracted rays are considered. Since the geometrical theory is simple to use and requires no integration, it seems to be preferable to the Braunbek method. It can also take account of corners and of multiple diffraction. On the other hand, the geometrical theory involves supplementary considerations in order to describe the field on caustics and shadow boundaries. The Braunbek method encounters no difficulty on such surfaces. Therefore it can be used to supplement the geometrical theory.

We also see why the simpler Kirchhoff methods predict the correct qualitative behavior of the diffracted field. This is because the Kirchhoff methods yield the same singly diffracted rays and the same phase as do the more accurate geometrical theory and the Braunbek method. Since the qualitative properties of the diffraction pattern depend upon interference between the fields on these rays, they are predicted essentially correctly. Quantitatively the Kirchhoff methods are also fairly accurate in and near the forward direction. This is because the Kirchhoff diffraction coefficients are close to the more correct coefficients of the Braunbek method and of the geometrical theory of diffraction in and near the forward direction. However away from this direction the Kirchhoff methods must be quantitatively inaccurate. They will also fail at smaller values of k when multiple diffraction becomes important.

Finally we have seen that the Kirchhoff method yields no correction, and the Braunbeck method yields an improper correction, to the geometrical optics transmission cross section of an aperture when it is computed by means of the cross section theorem. For a screen on which $u = 0$, this is due to the failure of these methods to take account of multiple diffraction. In the geometrical theory of diffraction the appropriate correction to the cross section for a circular hole in such a screen is obtained as soon as doubly diffracted rays are taken into account.

Appendix I. Interpretation of u^I

We have assumed that the incident field is asymptotically (for large k) of the form $Ae^{ik\Psi}$. It then follows[10] from the wave equation that Ψ satisfies the eiconal equation $(\nabla\Psi)^2 = 1$ and that A is given by

$$(A1) \qquad A(P) = \frac{A_o}{\sqrt{(1+\rho_1^{-1}R)(1+\rho_2^{-1}R)}} \quad .$$

In (A1) A_o denotes the value of A at some point P_o on a ray, A(P) is the value of A at P, a distance R along the ray from P_o, and ρ_1 and ρ_2 are the principal radii of curvature of the wavefront Ψ = const. at P_o. The result (A1) also follows from the law of conservation of energy within a tube of rays. From the eiconal equation we find that $\Psi(P) = \Psi_o + R$.

We wish to show that each term in equation (15) for $u_1^I(P)$ and $u_2^I(P)$ is of the form $A(P)e^{ik\Psi(P)}$ where A(P) and $\Psi(P)$ are determined as indicated above. It will then follow that both u_1^I (P) and $u_2^I(P)$ equal $u_{inc}(P)$, if P is in the lit region. From (15) we see at once that the phase has the desired form and that the amplitude is proportional to the value of A at the stationary point. Therefore calling this point P_o, we may identify $A(P_o)$ with A_o in equation (A1) above. It remains to be shown that the coefficient of $A(P_o)$ in (15) equals that of A_o in equation (A1).

The factors in the denominator of (A1) may be negative if the radii ρ_1 and/or ρ_2 are negative. If one factor is negative the root to be taken is $e^{i\pi/2}$ multiplied by a positive real number. If both factors are negative the root to be taken is $e^{i\pi}$ multiplied by a positive real number. These choices correspond to the fact that the phase on a ray changes by $-\frac{\pi}{2}$ each time the ray goes through a caustic[11] and when this occurs a factor of the denominator becomes negative. Therefore we can rewrite (A1) by introducing ε' which is +1, -1, or -i according as neither, both, or one center of curvature of the wavefront lies between P_o and P. If we also introduce the mean curvature $M = \rho_1^{-1} + \rho_2^{-1}$

and the Gaussian curvature $C = \rho_1^{-1}\rho_2^{-1}$ of the wavefront, (A1) becomes

$$\text{(A2)} \qquad A(P) = \frac{\varepsilon' A_o}{\sqrt{|1+MR+CR^2|}} \, .$$

This must be compared with the coefficient of $A(P_o)$ in (15), which is $\varepsilon \sqrt{(1-\Psi_x^2 - \Psi_y^2)/|\triangle|}$. To evaluate $\triangle$ we note that $R_{xx} = \left(\frac{x-x_o}{R}\right)_x = \frac{(1-R_x^2)}{R} = \frac{1-\Psi_x^2}{R}$, with similar expressions for R_{yy} and R_{zz}. Inserting these expressions into the definition of $\triangle$, equation (16), the coefficient of $A(P_o)$ becomes

$$\text{(A3)} \qquad \varepsilon \left[\left[R^2 \frac{\Psi_{xx}\Psi_{yy}-\Psi_{zz}}{1-\Psi_x^2-\Psi_y^2} + R\, \frac{(1-\Psi_y^2)\Psi_{xx}+2\Psi_x\Psi_y\Psi_{xy}+(1-\Psi_x^2)\Psi_{yy}}{1-\Psi_x^2-\Psi_y^2} + 1\right]\right]^{-1/2} .$$

It is now easy to show that the coefficient of R^2 in (A3) is just the Gaussian curvature C and the coefficient of R is the mean curvature M of the wavefront at P_o. It also follows from their definitions that ε and ε' are equal. Thus the coefficient of A_o in equation (A2) is equal to the coefficient of $A(P_o)$ in (15) which is given by equation (A3). This completes the proof of the fact that both $u_1^I(P)$ and $u_2^I(P)$ are equal to $u_{inc}(P)$, if P is in the lit region.

Appendix II. The geometrical factor G

In order to simplify the expression for G given in (22) we must evaluate $(\Psi+R)_{\sigma\sigma}$ at an edge stationary point. To do so it is convenient to introduce t, the unit vector tangent to the edge. Then $(\Psi+R)_\sigma = \nabla\Psi\cdot t + \nabla R\cdot t = \cos\beta + \nabla R\cdot t$. Therefore $(\Psi+R)_{\sigma\sigma} = -\beta_\sigma \sin\beta + \nabla R\cdot t_\sigma + (\nabla R)_\sigma\cdot t = -\beta_\sigma \sin\beta + \rho^{-1}\nabla R\cdot n + R^{-1}\sin^2\beta$. Here we have substituted $t_\sigma = \rho^{-1}n$ where ρ is the radius of curvature of the edge and n the unit normal. In showing that $(\nabla R)_\sigma\cdot t = R^{-1}\sin^2\beta$ we made use of the fact that $\nabla R\cdot t = -\cos\beta$ at the stationary point and we performed a simple calculation. If we now define the angle δ by $\cos\delta = -\nabla R\cdot n$ then (22) becomes

$$(A4) \qquad G = \frac{\sin\beta}{R\sqrt{\left|-\beta_\sigma \sin\beta - \rho^{-1}\cos\delta + R^{-1}\sin^2\beta\right|}}\, e^{-i\pi/4 \pm i\pi/4} .$$

Let us now define ρ_1 by

$$(A5) \qquad \rho_1 = \frac{-\sin^2\beta}{\beta_\sigma \sin\beta + \rho^{-1}\cos\delta} .$$

Then (A4) becomes

$$(A6) \qquad G = \left[\left|R(1+\rho_1^{-1}R\right|\right]^{-1/2} e^{-i\pi/4 \pm i\pi/4} .$$

If we omit the exponential factor and the absolute value signs under the radical, G is unchanged provided that the square root is determined in the manner explained just preceding equation (A2). This is obviously the case if the radicand is positive, while if it is negative the root introduces the necessary $e^{-i\pi/2}$. Thus we obtain equation (26), as we desired to show. It is to be noted that the definition of ρ_1 above in (A5) is exactly the same as that in equation (11) of Part I.

Appendix III. Asymptotic evaluation of Rubinowicz'line integral

Suppose that the incident wave is the spherical wave e^{ikr}/r, where r denotes distance from the source. A. Rubinowicz[4] has shown that then u_K can be written in the form

$$(A7) \qquad u_K = u_{geom} - \frac{1}{4\pi} \int \frac{e^{ik(r+R)}}{rR} \; \frac{\sin \beta \; m \cdot \nabla R}{1 + \nabla r \cdot \nabla R} \, d\sigma \; .$$

The integral is a line integral along the edge of the aperture and the vector m is the unit outward normal to the cone formed by lines joining the source to the points of the edge. Rubinowicz has evaluated this integral by the method of stationary phase. We find that this procedure applies to edge stationary points and yields

$$(A8) \qquad u_K^{II} \sim \sum D_k r^{-1} e^{ik(r+R)} G \; .$$

Here the sum extends over the edge stationary points, which are determined by the condition $(r+R)_\sigma = 0$. These are just the same stationary points as those obtained in evaluating (20). The factor G is given by equation (22) and D_k is defined by

$$(A9) \qquad D_K = - \frac{e^{i\pi/4} \; m \cdot \nabla R}{2 \sqrt{2\pi k} \; (1 + \nabla r \cdot \nabla R)} \; .$$

From its definition, we see that m is perpendicular to both ∇r and t. Since it is a unit vector, we have $m = t x \nabla r / \sin \beta$. Then $m \cdot \nabla R = t x \nabla r \cdot \nabla R / \sin \beta$. In order to evaluate this triple product it is convenient to consider the normal vector n, pointing into the aperture, and the binormal b, pointing toward the dark side. Then we have

$$\text{(A10)}\quad \nabla r = t \cos \beta - n \sin \beta \sin \alpha + b \sin \beta \cos \alpha ,$$

$$\text{(A11)}\quad \nabla R = -t \cos \beta - n \sin \beta \sin \theta + b \sin \beta \cos \theta .$$

Upon evaluating the triple product and the scalar product $\nabla r \cdot \nabla R$, and inserting the results into (A9) we obtain

$$\text{(A12)}\quad D_K = + \frac{e^{i\pi/4}(\sin \alpha \cos \theta - \cos \alpha \sin \theta)}{2 \sqrt{2\pi k} \sin \beta(1+\sin \alpha \sin \theta+\cos \alpha \cos \theta)} .$$

After some trigonometric transformations we find that D_K given by (A12) is exactly the same as D_K given by (29). Therefore the stationary points of the line integral contribute exactly the same terms to the field u_K as do the edge stationary points of the double integral (6), which also represents u_K.

Let us now consider the contributions of corners to the line integral. These contributions can be obtained by integrating by parts, and ignoring the remaining integral which is of lower order in k. In this way we obtain

$$\text{(A13)}\quad u_K^{III} \sim \sum - \frac{1}{4\pi} \frac{e^{ik(r+R)}}{rRik(1+\nabla r \cdot \nabla R)} \left[\frac{\sin \beta \; m \cdot \nabla R}{(r+R)_\sigma} \right] .$$

The brackets indicate the jump or discontinuity of the enclosed quantity at a corner, and the sum extends over all corners. This sum can be rewritten in the form

$$\text{(A14)}\quad u_K^{III} \sim \sum C_K \cdot \frac{1}{r} \cdot \frac{e^{ik(r+R)}}{R} ,$$

where

$$(A15) \qquad C_K = \frac{i}{4\pi k(1+\nabla r\cdot\nabla R)} \left[\frac{\sin\beta \; m\cdot\nabla R}{(r+R)_\sigma} \right] .$$

Upon inserting the angles previously defined, we find that (A15) becomes exactly (36).

References

[1] Keller, J.B. - Diffraction by an aperture I; New York University, Institute of Mathematical Sciences, Research Report EM-92, 1956.

[2] Bouwkamp, C.J.- Diffraction theory; Reports on Progress in Physics, 17, 35-100 (1954). Also New York University, Institute of Mathematical Sciences, Division of Electromagnetic Research, Research Report EM-50, 1953.

[3] Braunbek, W. - Neue Näherungsmethode für die Beugung am ebenen Schirm; Zeits. für Physik, 127, 381 (1950).

- Zur Beugung an der Kreisscheibe; ibid., 127, 405 (1950).

[4] Rubinowicz, A.- The diffraction waves in Kirchhoff's theory of diffraction phenomena; Ann.d.Phys., 53, 257 (1917).

- Kirchhoff's theory of diffraction, ibid., 73, 339 (1924).

[5] van Kampen, N.G.- An asymptotic treatment of diffraction problems; Physica, 14, 575 (1949).

[6] Focke, J. - Asymptotische Entwicklungen Mittels der Methode der Stationären Phase; Verh. der Sãch. Akad. der Wiss. zu Leipzig, 101, 1(1954).

[7] Chako, N. - Asymptotic expansions of double integrals occurring in diffraction theory; to be published, New York University, Institute of Mathematical Sciences, Division of Electromagnetic Research, Research Report EM-100 (1956).

[8] Erdelyi, A. - Asymptotic expansions; Dover Publishers (New York, 1956).

[9] Levine, H. - Diffraction by a circular aperture at high frequencies; Tech. Rep. 51, Appl. Math. and Stat. Lab., Stanford University, Stanford, California, June 1956.

[10] Keller, J.B., Lewis, R.M. and Seckler, B.D. - Asymptotic solution of some diffraction problems; Comm. Pure and Appl. Math., 9, 207 (1956). Also New York University, Institute of Mathematical Sciences, Division of Electromagnetic Research, Research Report No. EM-81, 1955.

[11] Kay, I. and Keller, J.B. - Asymptotic evaluation of the field at a caustic; Jour. App. Phys., 25, 876 (1954). Also New York University, Institute of Mathematical Sciences, Division of Electromagnetic Research, Research Report No. EM-55, 1953.

NYU
EM-
96

Keller

AUTHOR

Diffraction by an

TITLE

aperture. II.

NYU
EM-
96

Keller

Diffraction by an
aperture. II.

Date Due

MAY 7 '59			
MAY 23 '[illegible]			
JAN 26 '62			
MAY 23 '63			
MAR 3 '65			
DEC 2 '68			
JAN 30 1982			
JAN 31 1983			
JUN 4 1986			
JAN [illegible]			
FEB [illegible] 1986			
MAR 5 1986			
	PRINTED	IN U. S. A.	